BBC RD 1993/7

Research Department Report

DYNAMIC RANGE CONTROL OF AUDIO SIGNALS BY DIGITAL SIGNAL PROCESSING

N.H.C. Gilchrist, B.Sc., C.Eng., M.I.E.E.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION

BBC RD 1993/7

DYNAMIC RANGE CONTROL OF AUDIO SIGNALS BY DIGITAL SIGNAL PROCESSING

N.H.C. Gilchrist, B.Sc., C.Eng., M.I.E.E.

Summary

It is often necessary to reduce the dynamic range of musical programmes, particularly those comprising orchestral and choral music, for them to be received satisfactorily by listeners to conventional FM and AM broadcasts. With the arrival of DAB (Digital Audio Broadcasting) a much wider dynamic range will become available for radio broadcasting, although some listeners may prefer to have a signal with a reduced dynamic range.

This Report describes a digital processor developed by the BBC to control the dynamic range of musical programmes in a manner similar to that of a trained Studio Manager. It may be used prior to transmission in conventional broadcasting, replacing limiters or other compression equipment. In DAB, it offers the possibility of providing a dynamic range control signal to be sent to the receiver via an ancillary data channel, simultaneously with the uncompressed audio, giving the listener the option of the full dynamic range or a reduced dynamic range.

Issued under the Authority of

Ian Childs

Head of Research Department

Research Department, Engineering Division
BRITISH BROADCASTING CORPORATION

(S-21)

1993

DYNAMIC RANGE CONTROL OF AUDIO SIGNALS BY DIGITAL SIGNAL PROCESSING

N.H.C. Gilchrist, B.Sc., C.Eng., M.I.E.E.

DYNAMIC RANGE CONTROL OF AUDIO SIGNALS BY DIGITAL SIGNAL PROCESSING

N.H.C. Gilchrist, B.Sc., C.Eng., M.I.E.E.

1. INTRODUCTION

The wide dynamic range of some types of musical programme material (e.g. orchestral and choral music) necessitates the use of compression when this is to be broadcast by a conventional FM or AM transmitter. If compression is not used, there is the risk that either the signal peaks will over-modulate the transmitter, with consequent distortion of the sound, or that low-level passages will be adversely affected by noise. In the latter case, the broadcast signal sounds 'quiet' or 'under-modulated' when compared with compressed broadcast signals.

Broadcasters have used amplitude limiters for many years, not only to protect transmitters from over-modulation, but also to compress the loudest parts of the programme. Compression by a limiter invariably introduces audible impairment when used on music with a wide dynamic range, and it is generally recognised that the best results are obtained when a trained operator (a Studio Manager in radio, or a Sound Supervisor in television) controls the audio programme level, as is normally the case when 'live' concerts are broadcast.

With the introduction of Compact Disc (CD), recorded musical material with a wide dynamic range has become readily available to both the listener and the broadcaster. Material with a wide dynamic range may also be recorded on Digital Audio Tape (DAT) and Digital Compact Cassette (DCC). When broadcasting such wide-range material by FM and AM it would be possible to use a Studio Manager to effect the necessary compression, but experimental work is in progress at the BBC on the use of digital signal processors (DSP) to effect 'artistic' (i.e. unobtrusive) dynamic compression. The Studio Manager discreetly raises the level of quiet passages in the music, and anticipates an approaching *fortissimo* by slowly (and, hopefully, unobtrusively) reducing the level prior to the musical climax. The intention is to compress the dynamic range whilst preserving the impact of a dramatic change in the programme dynamics (e.g. that caused by a *crescendo* or *subito fortissimo*). A DSP has neither the Studio Manager's knowledge of music nor the ability to read a musical score, but it can gain experience of the programme content prior to the broadcast of a CD either by being able to examine the recording in advance or by delaying the programme in memory whilst making the necessary level adjustments.

The digital processing described in this Report uses the latter principle; that of delaying the programme in memory[1,2]. Because the processing does not require any prior knowledge of the material to be broadcast, it may be used on 'live' programmes.

2. THE COMPRESSION ALGORITHM

The principal characteristic of the compression algorithm* is a gain law, indicating the output level from the dynamic range controller as a function of the input level. Generally, this indicates that quiet signals are made louder, and loud signals quieter. For ease of operational use, the levels are specified in terms of the scale on a BBC Peak Programme Meter (PPM)[3]. Some examples of possible gain laws are shown in Figs. 1 to 3 *(overleaf)*. The law shown in Fig. 1 would provide limiting at both extremes of the dynamic range; the law in Fig. 2 would compress all dynamics; that in Fig. 3 combines both of the previous laws.

Ideally, as the level of the input signal varies, the changes in gain required to bring the output to the desired level would be made unnoticeably. To try to achieve this, the algorithm looks three seconds ahead into the audio data. It can thus anticipate impending changes in level and start to make gradual gain changes. The figure of three seconds was chosen as a compromise between better anticipation and ease of operational use. As the delay is made shorter there is less warning of the changes in level, so more abrupt changes in gain sometimes have to be made.

A schematic diagram of the controller is shown in Fig. 4 *(overleaf)*. The controller comprises a delay and a signal processor, with the undelayed samples providing the look-ahead for the signal processor. The delayed samples are multiplied by the appropriate gain values and then sent to the output.

The undelayed samples are analysed to provide a series of PPM values, one every quarter of a second. The value of gain (or attenuation) that would be needed to be applied to the input signal, to bring the peak in the 3-second time window to the desired level, is derived from the gain law. Although the gain law may be specified at only a few discrete points, the processor interpolates linearly between them to

* The algorithm has been developed by S.G. Tunnicliffe Wilson, A.K. McParland and A.J. Mason.

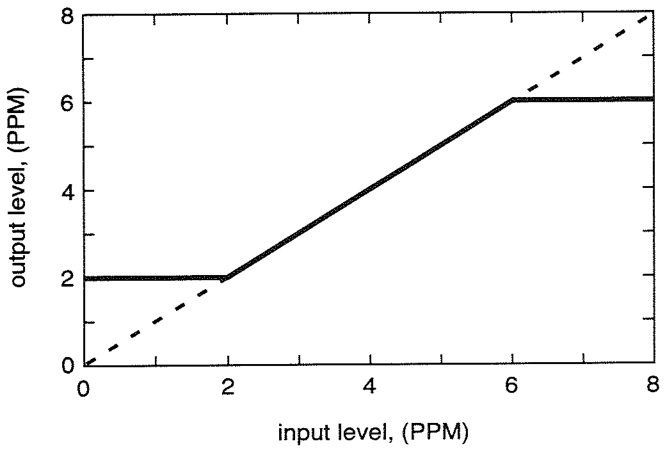

Fig. 1 - Law with limiting at extremes of level.

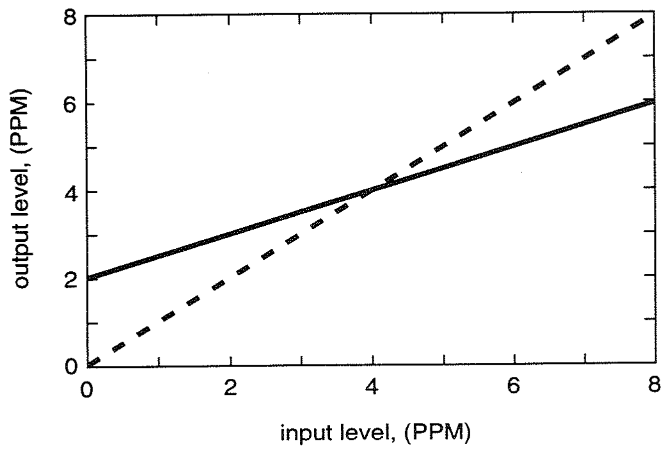

Fig. 2 - Law giving compression over the entire dynamic range.

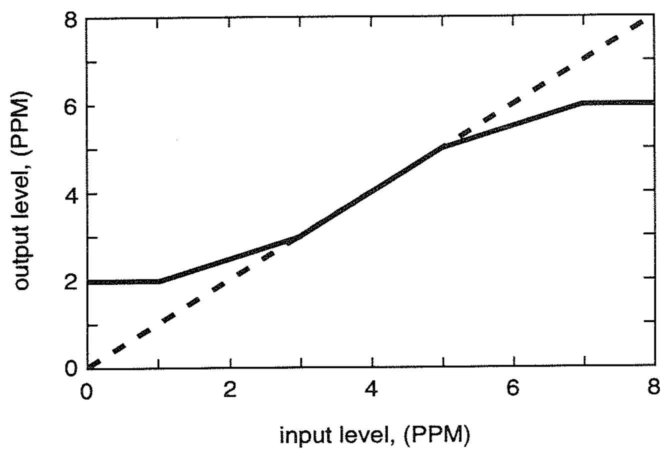

Fig. 3 - Law combining the characteristics of the laws in Figs. 1 and 2.

calculate the desired output signal level for any input signal level. The processor aims to have the gain at the required value by the time the peak in the window appears at the output of the delay; this will not always be achieved, however.

The maximum rate at which the gain can change unobtrusively to that required is preset in the program. A figure of just over 1 dB per second was used initially in the BBC's experimental implementation. Recent listening tests have shown that the increases in gain intended to raise the level of quiet signals were more obtrusive than the reductions. In one item of choral music the gradual increase in gain applied to a quiet passage gave the disturbing impression that the whole choir was creeping forwards. It was necessary to reduce the rate of gain increase to about 0.5 dB per second to overcome this effect. The maximum rate of decrease in gain remained at about 1 dB per second.

Once calculated, the gain values are smoothed so that they do not continuously follow the smallest variations in signal level.

The limit that is placed on the maximum rate of gain reduction is ignored if the output level would otherwise exceed a predetermined 'overload' level. This ensures that equipment further along the broadcast chain is not driven to the point where it may clip or limit the signal in a crude fashion. Such impairments as may be detected in the processed signals occur when this condition exists, but these tend to be much less annoying than the alternatives.

A limit is set to the maximum gain which can be applied to the very quietest signals. There are two reasons for this. The first is that quiet sources, such as soft singing, can sound unnatural when reproduced at an excessively high level. The second is that background noise may be raised to the level where it becomes obtrusive.

The algorithm implemented as described above works fairly well. However, some impairments were noticed on certain critical types of material. These led to the addition of some secondary characteristics to the algorithm.

One item of music which revealed unnatural sounding impairments was a passage of piano music containing a gradual *diminuendo*. As the level of the input signal fell, the gain applied was increased. In this case, a peculiar sustaining of the decay of individual piano notes was noticed, during the *diminuendo*. Each decay took fractionally longer to extinction than was expected, and after a short period of listening the effect became very noticeable. Similarly, when there were very small *crescendi* in an item of music in which the programme level was generally rising, the gradual reduction in gain which was occurring tended to obliterate them.

The secondary characteristics which were

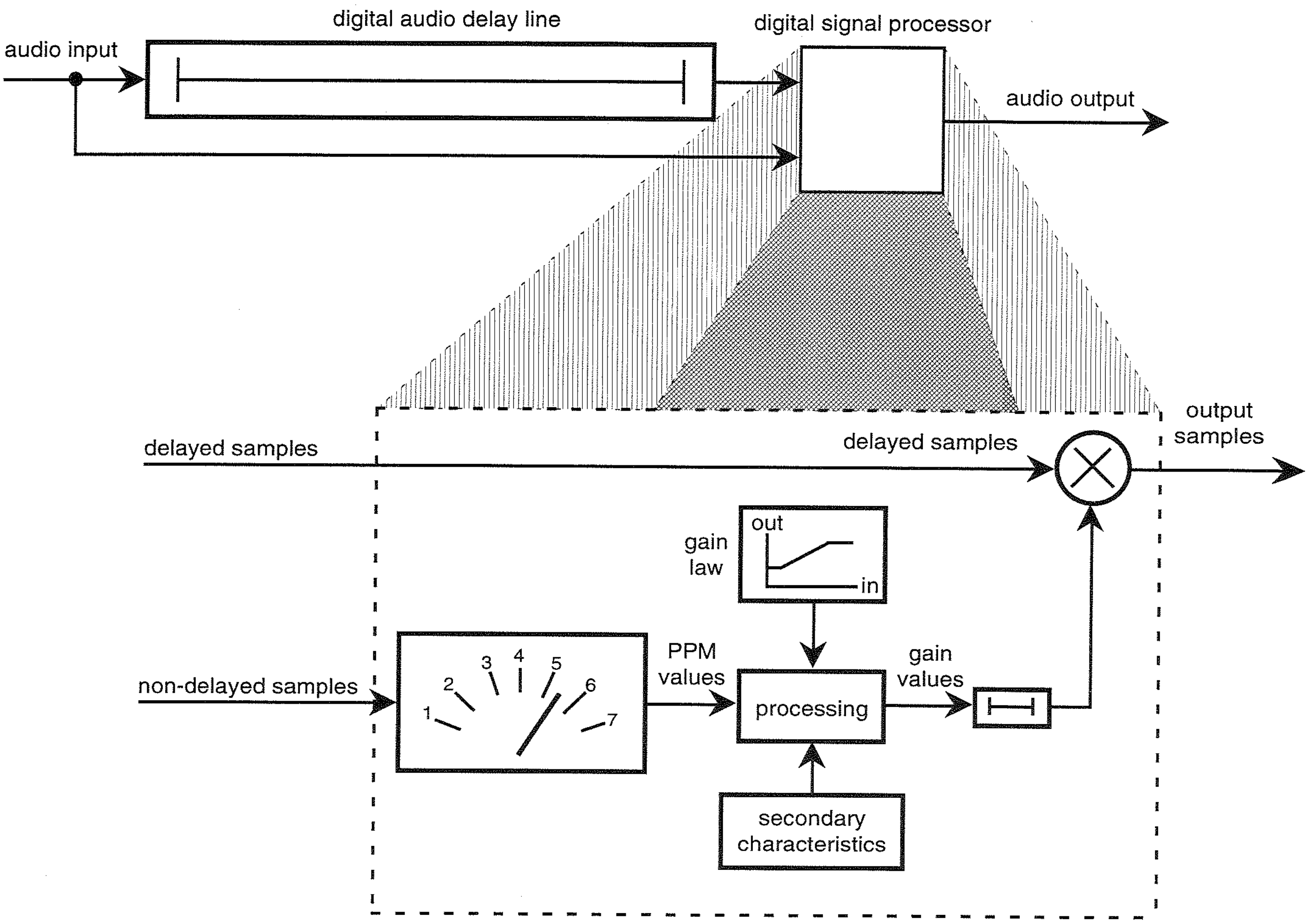

Fig. 4 - Block diagram of the dynamic range controller.

added were as follows. Whilst the gain is being increased, if two or more successive quarter-second blocks of samples show decreasing level, then the gain is not increased during those blocks. While the gain is being decreased, if two successive quarter-second blocks show increasing level, then the gain is not reduced during those blocks (although priority is given to ensuring that the output signal does not exceed the predetermined level at which downstream equipment would introduce limiting).

The inclusion of these secondary characteristics removes the peculiar artifacts. The most obvious effects of the processing which remain occur when a very loud passage follows a very quiet one. Under these circumstances, the dynamic range controller has to change from maximum gain to maximum attenuation in the 3-second look-ahead time, to prevent the signal peaks exceeding the limiting level of equipment later in the broadcasting chain. To do this it has to override the limit on the maximum rate of gain change.

3. NON REAL-TIME PROCESSING

The software to implement the dynamic range controller was originally written for a Sun workstation, not running in real time. An AES/EBU interface was developed for the workstation, to give real-time exchange of digital audio between the hard disk storage and external sources and destinations, for recording and replay of test programme material. Such an arrangement has the advantages of ease of change and good debugging facilities. The disadvantage is that the compression software does not run in real time. Thus the test programme material could only be heard in its compressed form after a period of time spent in recording and processing.

The compression software was written in the 'C' language, which facilitated fast algorithm development. The algorithm was developed in this environment to the stage where good results were obtained with a wide variety of signals.

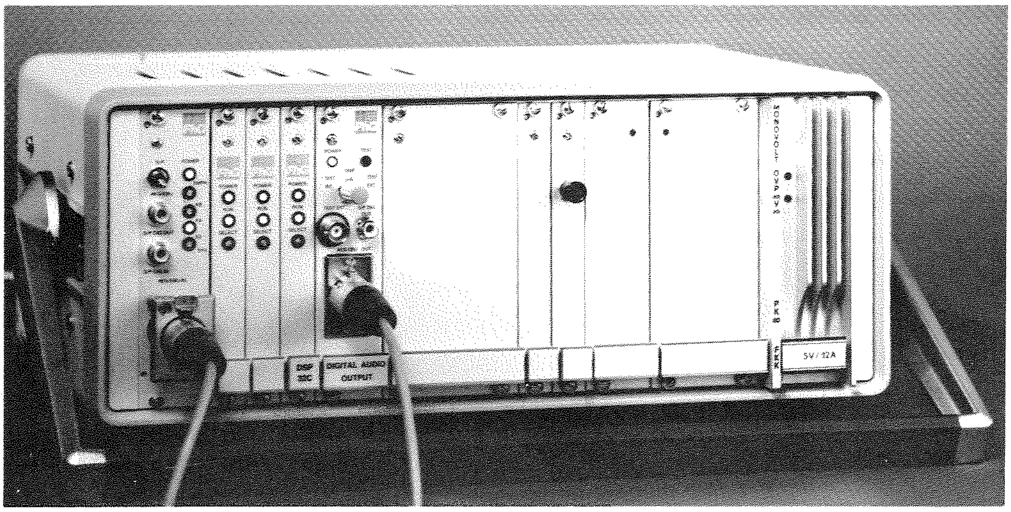

Fig. 5 - The real-time hardware dynamic range controller.

4. REAL-TIME HARDWARE

When the development of the dynamic range control processor had reached the point where promising results could be demonstrated, the change to real-time implementation was justified. With the algorithm running in real time, the further development of the algorithm could take place more rapidly, because changes could be evaluated immediately. Another benefit was that a practical item of hardware could be demonstrated to operational colleagues, and consultation could take place to enable the characteristics to be optimised prior to any in-service trials.

4.1 The dynamic range controller

General-purpose audio digital processing hardware was available, with a good 'C' compiler and some development utilities. The dynamic range compression software was transferred from the Sun workstation to this hardware to provide the real-time implementation.

The real-time hardware* is shown in Fig. 5. It comprises a 3U Eurocard rack with AES/EBU input and output interface units and a number of digital signal processor (DSP) units. The DSP used is the DSP32C floating-point processor which is capable of operating at speeds up to 25 Mflops (25 million floating-point operations per second). This particular system is very flexible, as each processor unit can pass on data to the next as a serial bitstream in the form of a user-defined 'C' structure. Thus, if the algorithm is too demanding for a single processor unit, the task can readily be shared between a number of such units. 'Wrapper' programs, written in 'C', are provided with the hardware, and the designer's code is placed within these.

Three processors were used in the implementation shown in Fig. 5, however, this is not an indication of the processing power required. The reason for using the three processor units is simply to provide the three seconds' delay needed for the look-ahead time mentioned earlier. Each processor unit has sufficient RAM available (256 kbytes) for just over one second's delay, at 44.1 kHz sampling frequency, plus the processing code. In this implementation, the first processor does all the calculations and passes the calculated gain through the second processor to the third. The third processor then applies the required gain changes to the delayed audio samples.

Any item of programme may be reviewed using the real-time compressor. The original software has been extended to enable many of the parameters to be changed, so that different gain laws, rates of gain change and decision thresholds may be tried out. This has facilitated further refinement of the compression technique, to make the compression more and more subtle (i.e. less obtrusive).

4.2 The user interface

In order to enable operators and experimenters to adjust the parameters of the dynamic compressor with ease, controlling software was developed to run on a laptop PC (Personal Computer)**. The PC is connected to the dynamic range controller via an RS 232 serial interface, and runs software which has been written using Turbo C++ and Turbovision.

The user is presented with a screen display on the PC comprising three windows: a graph, a meter display and a status window. These are shown in Fig. 6. The meter display, at the foot of the screen,

* The processing hardware was made by the Fraunhofer Institute for Integrated Circuits, Erlangen, Germany.

** The controlling software was developed by J.G. Walker.

gives input and output stereo PPM level readings in the form of bar-graph indicators and also as numerical values. It also indicates the gain, expressed in dB, being applied to the audio signal. The status window at the top right-hand side shows the look-ahead delay being used, and indicates which of the facilities listed are being used. At the top left-hand side, the graph window displays the compression law which is in use.

At the very top of the screen in Fig. 6 is a menu bar. The user has the option to change parts of the display or some of the parameters of the compressor, using the menu. On accessing the 'DSP Communications' item, the user can arrange to inspect DSP memory locations, display characters and numerical values, edit the DSP address list, enable or disable diagnostic windows and open or close the status window. The programme level meter may be PPMs or simple absolute digital level meters, selected via the 'Meters' item, or the meter window may be closed. The 'Compression Law' item enables the user to open and close the graph window, and to edit the input/output PPM characteristics. Finally the 'Options' item enables the user to change the type of compression, change the parameters of the compression mode selected, save a particular version of the compressor, recall a previously-saved compressor, change the screen display colours, by-pass the compression process or make a mono mix of the stereo.

Fig. 7 *(overleaf)* shows the screen display with the 'Options' dialog box open. A long-delay compression process is selected, and various processing options and parameter values are displayed.

5. DIGITAL AUDIO BROADCASTING

With the prospect of having digital audio broadcasting (DAB) within the next few years, a broadcasting medium will be available which is capable of carrying programme material with the dynamic range of CD, DAT and DCC recordings without compression. No longer will the situation exist in which the dynamic range has to be influenced by the characteristics of the broadcasting medium.

The introduction of DAB will not remove the need for dynamic compressors, but it is likely to change the way in which they are used. Some listeners will be in a position to appreciate having the wider dynamic range which is available, but others may prefer to listen to a compressed programme. Listeners in noisy environments, including those travelling in cars, are likely to benefit from a dynamically compressed signal. The broadcaster should, therefore, provide the listener with a broadcast signal having the full dynamic range, together with a control signal derived from a compression processor. The control signal could then be used, optionally, in the receiver to

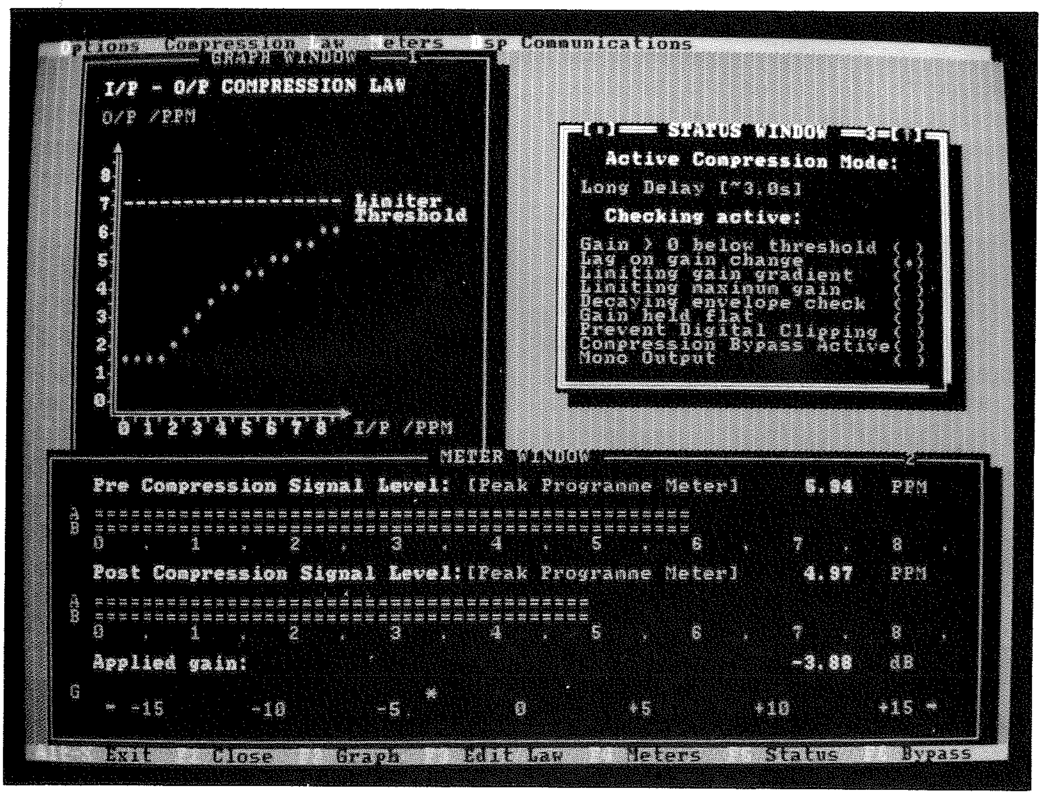

Fig. 6 - The user interface screen display with graph and status window.

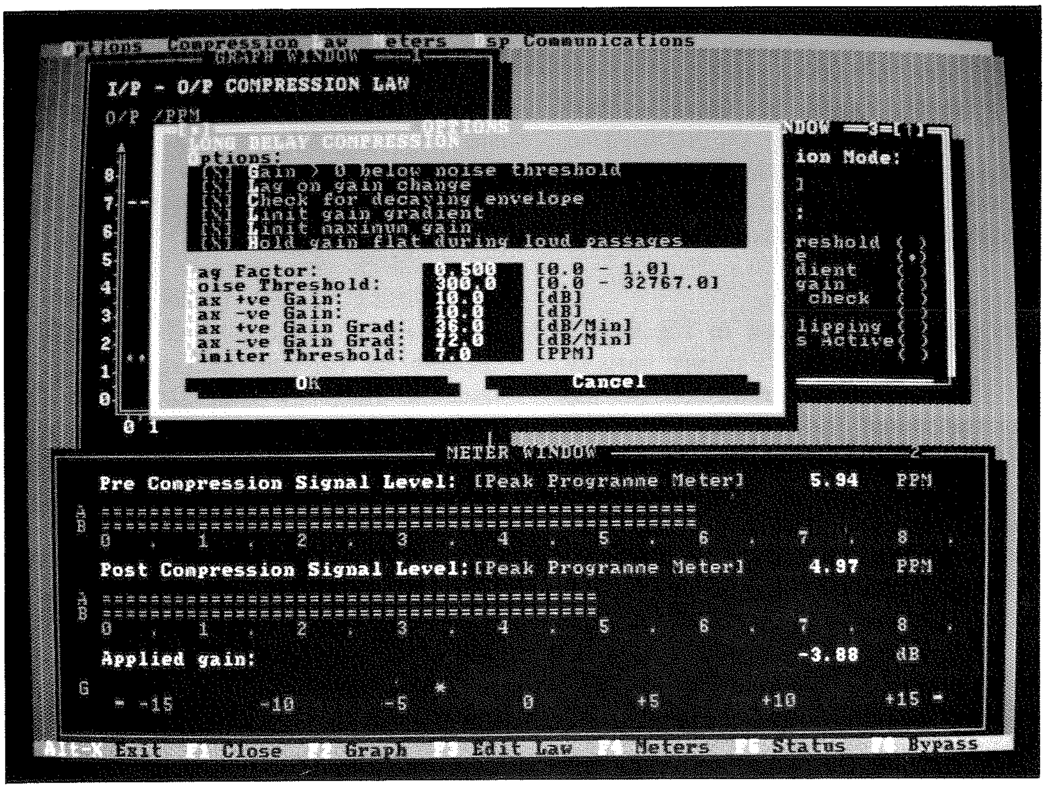

Fig. 7 - The user interface screen display with the 'options' dialog box open.

control a variable-gain process to effect the compression. If the broadcaster derives the control signal using the type of processing described in this Report, the compression applied will be relatively unobtrusive. There is the possibility that the listener could adjust the degree of compression introduced, according to personal taste, by modifying the control signal in the receiver.

In the DAB system, the delayed, but as yet uncompressed, audio samples would be taken from the output of the digital audio delay line in the compression processor (see Fig. 4) for onward transmission. The gain values from the processing (in the lower path through the processor, shown in the lower part of the diagram of Fig. 4) would be encoded in the dynamic range control data (DRCD) and sent as programme-associated data in the DAB signal. The Eureka 147 Digital Audio Broadcasting project has provisionally allocated 6 bits in one of the bytes of the PAD (Programme-Associated Data) for the use of the DRCD.

6. IMPLICATIONS FOR THE AES/EBU INTERFACE

The generation of additional data within a broadcasting centre always raises the question as to how the data may be conveyed with the audio signal. In the digital environment, audio signals are exchanged by means of the AES/EBU digital audio interface[4 - 6]. This interface can carry additional data, as user data or as channel status information. The format of a single sub-frame of the AES/EBU interface signal is shown in Fig. 8. The user and channel status bits follow the audio sample word, at the end of the sub-frame, and there are two sub-frames per frame; one for the right channel and one for the left channel of a stereo pair.

There is, however, a problem with putting the DRCD into the user or channel status bits. Each time that a signal is re-framed, the DRCD would have to be delayed in order to await the appropriate point in the new frame for its insertion into the bitstream. Fortunately, the compression processor is likely to be located in the continuity area of a broadcasting centre, and this is the point at which there is no longer the requirement to carry co-ordination signals in the four auxiliary sample bits of the sub-frame (see Fig. 8)[5, 6]. By taking the auxiliary sample bits for both sub-frames of each frame together, one byte per left/right sample pair is available for the 6-bit DRCD. The signalling channel provided by the auxiliary sample bits is ideal for this purpose, because the data refresh rate and sample rate are one and the same. Operations such as sampling-frequency changing, sampling-frequency synchronisation and (finally) reformatting the audio and DRCD into the DAB frame structure will always

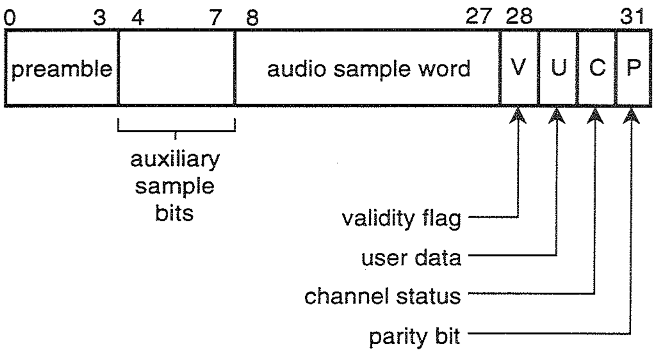

Fig. 8 - The sub-frame format of the AES/EBU interface signal.

be provided with the DRCD co-timed with the audio samples. Thus the possibility of introducing cumulative differential delay between the audio and the DRCD may be avoided.

7. OPERATIONAL CONSIDERATIONS

The DSP-based dynamic compressor needs to introduce a delay of several seconds into the programme, and for this reason careful consideration must be given to its use in broadcasting operations.

Delays of even a few milliseconds in distribution or contribution can interfere with the off-air cueing which is sometimes used in present-day broadcasting[7]. There is the possibility that operational practices may change to accommodate delays in broadcasters' connections. At one time it was unthinkable for some broadcasters to operate with more than a few milliseconds' delay in a connection. However, with satellite connections and digital systems being used more and more, there is some evidence that broadcasters are learning to live with the consequences of delay in some of their operations.

The compressor was originally intended for use with CD or DAT replay equipment. The delay between starting replay and the emergence of the signal from the compressor might be somewhat unnerving for an operator, but it is probably not likely to be too problematic for such programmes (e.g. choral and orchestral music) where it is normal to allow a pause between an introductory announcement and the replaying of a recording. In this particular application, if a delay in start-up were to cause difficulties, it should be possible to eliminate this by arranging for the replay equipment (CD or DAT) to fill the buffer memory of the dynamic compressor, and then halt in the 'pause' mode until the replay of the programme is cued to start.

In post-production work, the introduction of delays into connections between recording and replaying equipment are unlikely to cause problems.

Once a DAB system is implemented, if there is the need to duplicate programmes on FM (or vice versa) the FM transmitters could be fed from DAB signals received off-air, with dynamic compression applied under the control of the DRCD. AM transmitters could also be fed from DAB in a similar way, but possibly with a greater degree of compression applied.

8. CONCLUSIONS

A digital audio compression algorithm has been developed, which may be used to reduce the dynamic range of audio signals in an unobtrusive, or 'artistic', manner. The algorithm runs in real time in experimental hardware, and may be used either in a self-contained compressor (e.g. for feeding conventional broadcast transmitters) or to derive a compression control signal for applications such as DAB, where the compression is to be applied at a later point in the broadcasting chain.

A method for conveying the compression control signal through the AES/EBU digital audio interface has been proposed.

9. REFERENCES

1. BBC, 1992. Dynamic range compression. Inventor: TUNNICLIFFE WILSON, S.G. UK Patent No. 92 06065.6.

2. BBC, 1992. Dynamic range compression. Inventor: TUNNICLIFFE WILSON, S.G. International Patent No. PCT/GB92/00502.

3. CCIR, 1990. Recommendation 645. Test signals to be used on international sound-programme connections. Seventeenth Plenary Assembly (Düsseldorf, 1990), Green Book, **Vol. X**, Part 1, pp. 174-176.

4. EBU, 1985. Specification of the digital audio interface. EBU Publication No. Tech. 3250-E.

5. AUDIO ENGINEERING SOCIETY, 1991. Draft AES recommended practice for digital audio engineering — serial transmission format for two-channel linearly represented digital audio data. Revision of AES3-1985 (ANSI S4.40-1985).

6. CCIR, 1990. Recommendation 647. A digital audio interface for broadcasting studios. Seventeenth Plenary Assembly (Düsseldorf, 1990), Green Book, **Vol. X**, Part 1, pp. 182-195.

7. GILCHRIST, N.H.C. 1991. Delay in broadcasting operations. 90th AES Convention (Paris, 1991). AES Preprint No. 3033.

Printed by BBC RESEARCH DEPARTMENT, Kingswood Warren, Tadworth, Surrey, KT20 6NP